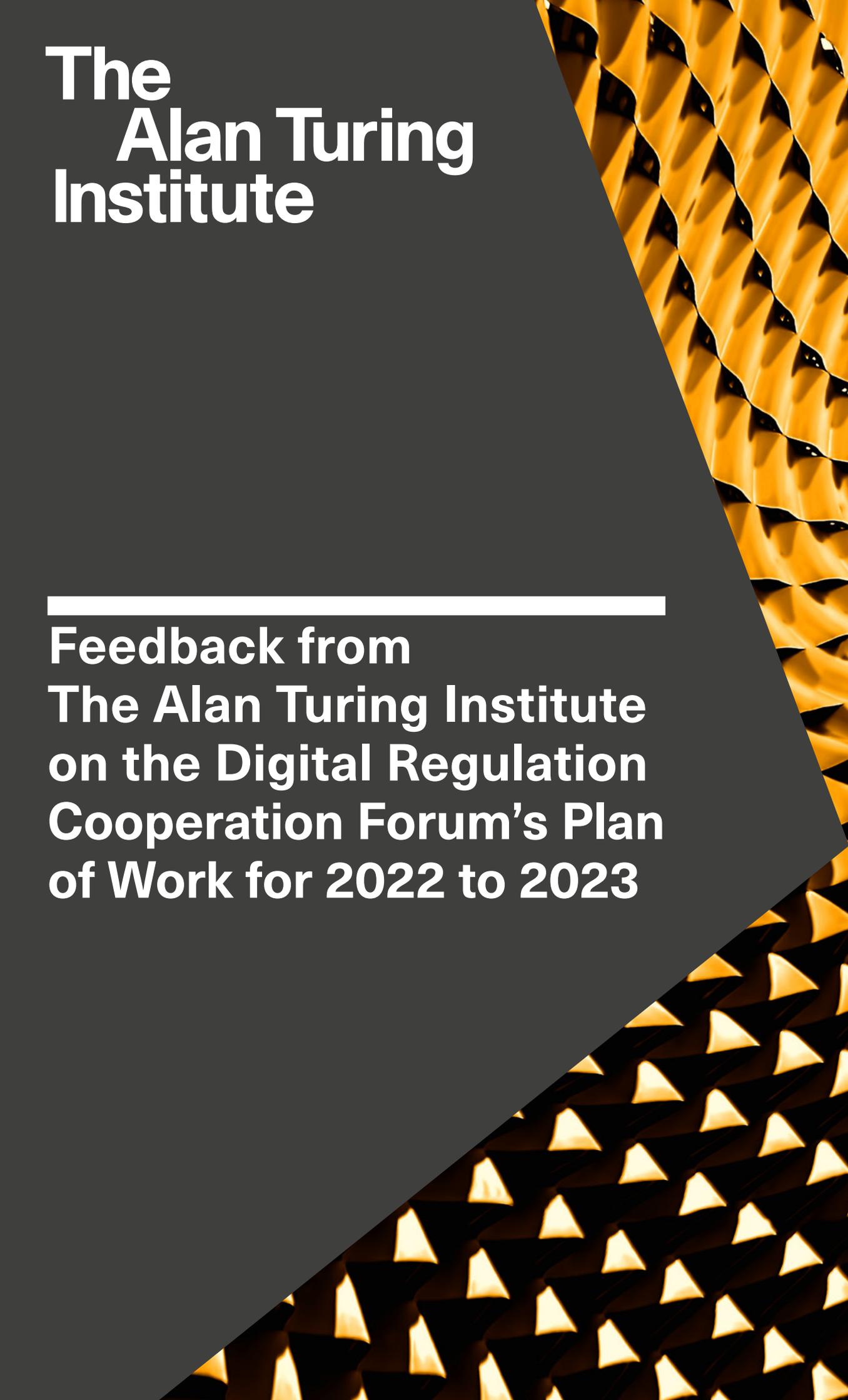


The Alan Turing Institute



**Feedback from
The Alan Turing Institute
on the Digital Regulation
Cooperation Forum's Plan
of Work for 2022 to 2023**



Content

The Alan Turing Institute	3
The Public Policy Programme	3
Executive summary	4
Rationale for submission	5
1. Coherence	6
2. Collaboration	9
3. Capability	12
4. Closing comments	16
5. Contributors	16

This document outlines **The Alan Turing Institute**'s feedback on the Digital Regulation Cooperation Forum (DRCF)'s ***Plan of work for 2022 to 2023***. The response synthesizes the perspectives of researchers at the Turing with expertise and interest in the area of AI regulation. The document's final section lists the names of the researchers who contributed to this response.

About The Alan Turing Institute

The Alan Turing Institute ('Turing') is the UK's national institute for data science and artificial intelligence (AI). Our mission is to make great leaps in data science and AI research in order to change the world for the better.

The Turing was created as a joint venture in 2015 between five leading UK universities – Cambridge, Edinburgh, Oxford, University College London and Warwick - and the Engineering and Physical Sciences Research Council (EPSRC). Subsequently, eight universities joined as partners - Birmingham, Bristol, Exeter, Leeds, Manchester, Newcastle, Queen Mary's, and Southampton. Today, we are a home to a community of more than 500 researchers, a growing team of in-house research software engineers and data scientists, and a professional services team. We convene activity across the data science and AI community through partnerships with universities and research institutes across the UK, as well as with industry, third sector and public organisations, with nine strategic partnerships.

About the Public Policy Programme

The Alan Turing Institute's **Public Policy Programme** works alongside policy-makers to explore how data-driven public service provision and policy innovation might help solve long-running societal problems. The programme also works hand-in-hand with regulators to develop practice-based guidance and standards for the responsible development and use of data science and AI.

Executive summary

The Turing sees the existence of collaborative regulatory initiatives like the DRCF as essential to the UK's ambition to be a global leader in the regulation of AI. Developing an effective and inter-organisationally robust regulatory environment is crucial for mitigating the far-reaching risks associated with AI. We welcome the DRCF's plan of work for 2022 to 2023 and we value its focus on the overarching goals of coherence, collaboration, and building capability. Informed principally by our Public Policy Programme's research, we offer the following recommendations on the plan.

Coherence

- In addition to the indisputable needs identified in the report, we recommend activities such as the development of a shared vocabulary to provide conceptual clarity for AI regulation, and comprehensive mapping of the risks that AI technologies pose across regulatory remits.
- For protecting children online, we highlight the need to develop an evidence base and improve understanding of age verification, including the risks associated with outsourcing to third parties and using biometric data.
- We recommend prioritisation of research into the relationship between online engagement and children's development, including the development of ways to involve children in such research.
- We recommend that DRCF consider incorporation of the principles from UNICEF's Policy Guidance on AI for Children
- Regulation of online advertising that negotiates the trade-off between individual rights and public benefit will require the enforcement of powers with respect to data sources; the pooling of expertise across regulators, and inclusive and participatory forms of democratic deliberation.

Collaboration

- The Turing has deep pockets of expertise and foundational work in algorithmic transparency. We are keen to support DCRF's work in this area
- We recommend a nuanced, interdisciplinary and practice-based approach to transparency that considers the transparency of processes, institutions and outcomes. We provide examples of Turing research that can inform this approach.

Capability

- The Turing has conducted research to understand the capabilities of UK regulatory bodies in relation to AI, both in terms of regulation of AI and the use of AI technologies as regulatory tools.
- We recommend that DRCF:
- build regulatory capability by developing skills across the regulatory landscape: fostering senior management understanding and buy-in; making training, skills development and tools across and beyond DRCF's constituent regulators.
- consider the rapidly evolving field of AI standardisation as an important area of capability building. The Turing Institute has been selected to lead the pilot of the **AI Standards Hub** set out in the Government's **National AI Strategy** and we would be delighted to collaborate with the DRCF with a view to enabling capability building in this area.
- encourage the broadening of stakeholder engagement, particularly citizens, through the incorporation of experiential expertise (taking account of end-user experience) in capacity building.

Rationale for submission

AI systems are growing in use and complexity. They are becoming a prevalent feature of products and services that individuals and organisations use daily, from customer chatbots and home smart speakers to fraud detection systems and CV screening tools.

As the prevalence of AI technologies increases, the remits of regulatory bodies will need to encompass the growing scope of AI applications and their impacts. The Digital Regulation Cooperation Forum directly addresses this need, underlining the importance of collaboration across regulators to identify and address the cross-sectoral challenges brought about by AI.

The Turing sees the existence of collaborative regulatory initiatives like the DRCF as essential to the UK's ambition to be a global leader in the regulation of AI. Developing an effective and inter-organisationally robust regulatory environment is crucial for mitigating the far-reaching risks associated with AI. Such an environment is needed to build confidence across industries and among innovators that the social and legal uncertainties brought about by AI's transformative force are being properly managed.

Within the UK, the DRCF is an important initiative intended to spur regulatory cooperation in online services. Thus, as the national institute for data science and AI, the Turing is committed to supporting the DRCF in its mission. The Institute's access to expertise in AI technologies, ethics, and regulation can act as a sounding board or support system for DRCF's ambitious initiatives in AI regulation. The Turing's mission is to make great leaps in data science and AI research in order to change the world for the better. Regulation is a vital part of turning this mission into reality. The Turing has a deep interest in ensuring that the network can contribute as much as possible to create a world-leading regulatory regime for AI technologies in the UK.

The Turing welcomes the publication of the DRCF's plan of work for 2022 to 2023 and its focus on the overarching goals of coherence, collaboration, and capacity building. The Turing's Public Policy Programme has conducted extensive research on what regulators need to do to adapt to a world that is increasingly turning to AI solutions. The results showed that readiness is the most important common denominator across the competencies required to foster effective regulation of AI. Readiness at a system-level – namely, the socio-economic, political, and inter-organisational circumstances and the general legal, regulatory, and policy surroundings that operate as preconditions of the successful adoption of new technologies – is closely connected to DRCF's goals of coherence, collaboration, and capacity building. These goals will be discussed in turn in the following sections.

1. Coherence

DRCF's goal of achieving coherence between regimes is crucial, and we welcome its inclusion in the Forum's plan of work for 2022 to 2023. Coherence between regimes provides clarity, consistency, and certainty across the regulatory landscape and gives regulated entities and the public the confidence needed to take advantage of technological innovation.

Some uses of AI technologies cut across regulatory remits. The Turing Public Policy Programme has interviewed employees from small, medium, and large size regulators within the UK to understand the need for common capacity for AI for the country's regulators. Several interviewees spoke of the challenges of regulating AI use cases that cut across traditional sectoral boundaries and emphasised the need for regulators to create consistent or complementary approaches.

When discussing coherence, the DRCF's plan of work identifies the need to spot "gaps and explore ways to apply regulations across different regimes," identify "areas of potential tensions" between regulatory regimes and work together to achieve regulatory objectives. Based on our research at the Turing, we would add some additional needs. Our interviews with regulators showed that coherence also involves:

- developing a shared vocabulary that provides conceptual clarity and a common language around AI;
- building a comprehensive mapping and understanding of the risks that AI technologies pose across regulatory remits;
- identifying the most efficient and effective ways of addressing regulatory gaps;, and
- anticipating and adapting to future risks and the speed of technological change.

DRCF's plan of work for 2022 to 2023 identifies key priorities under the overarching goal of coherence, which include protecting children online and promoting competition and privacy in online advertising. These are essential and much-needed priorities, and are discussed below.

1.1 Protecting children online

Children's needs in the digital age will require careful thought and coordination, so it is encouraging to see the DRCF turn its attention to this important topic. Turing research looking at public sector organisations' experiences in relation to AI and children's rights found that stakeholders were concerned that there are several guidance documents being drafted on the topic of children's rights and AI. Faced with competing guidance, organisations often feel unsure which to use moving forward and have indicated the need to see synergies formed between existing and upcoming guidance documents. There must be clear, accessible, and consistent guidance relating to children's rights and online safety to ensure compliance and to give regulated entities (as well as regulators) confidence in the approaches taken. The DRCF's proposed approach to bring together the work of Ofcom and the Information Commissioner's Office (ICO) is an encouraging step in the right direction.

Age assurance has benefited from renewed attention from UK legislators. The Online Safety Bill, for example, includes age verification for pornography websites to prevent underage use. DRCF's plan of work for 2022 to 2023 also mentions age assurance, which is reassuring. In addition to what is already in DRCF's plan, Turing would add the need to fully understand the risks associated with outsourcing age verification to third parties and the risks of age verification using biometric data and/or profiling.

More broadly, the Turing would encourage the development of a growing evidence base to understand parental and children's attitudes toward age assurance across different online contexts. The Turing would also welcome further research into how age verification and age estimation are but one piece of the age assurance puzzle. When viewing the path taken by the US through the Children's Online Privacy Protection Act (COPPA), it has become clear that establishing the age of consent is not a sufficient condition for protecting children online. There is an overarching question of individualised notions of consent versus the average child dilemma regarding consent and access to content. Should the age of consent for accessing certain online services be generalised, as it currently stands in GDPR, or should it be individualised to cater to different levels of maturity, development, and the unique needs of individual children? These questions need evidence-based answers.

Outside of the area of age assurance, research is needed to understand other aspects of how to consider the best interests of the child in a digital environment. Areas of concern include:

- the potentially transformative effects that online engagement has on children's development and their participation in the communities to which they belong
- the support that children might need after accessing a service that they are not supposed to be using at their age
- the need to consider the language that we are using of online safety vs empowerment and the desirability of shifting the focus from merely protecting children to educating, and
- equipping them with the capacity to make good decisions and become good digital citizens.

We recommend that the DRCF consider UNICEF's [***Policy Guidance on AI for Children***](#) when working towards the priorities of the plan of work for 2022 to 2023. This guidance contains nine principles for promoting child-centred AI, including "prioritising fairness and non-discrimination for children" as well as "preparing children for present and future developments in AI." This is especially pertinent, as several of these UNICEF's principles touch on elements not included in GDPR or the Children's Code.

The final recommendation in this area relates to the methods for generating the evidence-base needed to protect children online. [***Research at the Turing***](#) has found that public sector organisations do not involve children in conversations surrounding the design, development, and deployment of AI systems that use children's data. However, the Turing's interviews with public sector stakeholders found that they want to engage children on these topics, but they often do not know how to proceed. Guidance and support are needed in this area, as well as a growing body of research that engages with children to improve our understanding of their needs and preferences.

1.2 Promoting competition and privacy in online advertising

Competition and privacy in online advertising are complex issues to study. Within the UK, regulators like the Competition and Markets Authority (CMA) and the ICO have made significant progress over the past couple of years in understanding the benefits and risks of personalised advertising. It is encouraging to see them joining forces within the DRCF to tackle this issue jointly.

Important obstacles to understanding online platform advertising include lack of data, and the need for expertise across regulators. Researchers trying to understand, for example, the impact on democratic participation of personalised online advertising in UK elections need access to data held by social media platforms. It is widely known that algorithmic and data-driven techniques are used to infer sensitive information about social media users and target advertisements accordingly in ways that could impact electoral outcomes, yet

there is no way of identifying or quantifying such impact without data. Likewise, regulators concerned with elections need specific expertise in the fast-moving field of digital advertising and campaigning. The [House of Lords Select Committee on Democracy and Digital Technologies](#) has recently proposed a review of regulatory digital capacity across the CMA, ICO, Electoral Commission, Advertising Standards Authority and Ofcom to determine levels of the requisite digital expertise and where internal investment may be needed.

Similar technologies can also be used for many more forms of hyper-personalised and targeted messaging, as well as novel forms of personalised pricing. DRCF's constituent regulators have produced notable reports (for example, the CMA's [Pricing algorithms research, collusion and personalised pricing](#)), which bring to light the fact that these technologies can have discriminatory impacts on some of the most vulnerable users. Their use also makes it difficult to enforce compliance with existing regulation because the ex-ante duties or obligations for organisations to disclose relevant information are currently considered insufficient, and many ex-post methods of auditing cannot be enacted or enforced until substantial harm has been identified. As noted above, in the example related to elections, it can be challenging to prove such harm exists without access to the requisite information – potentially leading to a vicious cycle.

The knowledge acquired by individual regulators in online advertising, combined with their respective enforcement powers, should give a solid footing to the DRCF in pursuing this strand of work. As the DRCF is about to embark on this journey, it is worth noting that current proposals by organisations such as Google to remove third-party cookies or for Apple to implement additional privacy measures throughout their ecosystem complicate the picture further. On the one hand, such measures may benefit individual consumers by limiting the amount of data collected from their devices. However, such proposals could also further the market position of these companies, which may have benefited from potentially pervasive forms of data surveillance and extraction. While these initiatives may be touted as beneficial for consumers or complementary to individual rights (e.g. data protection and privacy), it may also hamper the ability of public organisations and regulators to operate effectively.

Determining how best to navigate the trade-off between individual rights and public benefit requires inclusive and participatory forms of democratic deliberation. At the Turing Institute, we have already encountered some success with this type of public engagement, such as our recent work with Camden Council to produce a [Data Charter](#) with 20 representative residents. The charter will guide how Camden Council collects, processes and shares data ethically and responsibly. Participatory processes such as this are vital for building trust and digital literacy among members of the public. As the DRCF turns its attention to promoting competition and privacy in online advertising, we would like to echo the importance of engaging with external stakeholders, including members of the public.

2. Collaboration

We welcome the inclusion of an overarching aim of collaboration in the DRCF's plan of work for 2022 to 2023. Regulators have their individual remits, but collaboration is a crucial enabler for effective oversight of general-purpose technologies that cut across sectors. It is encouraging to see regulators tackle an ambitious programme of work collaboratively. As they embark on this work, we would like to highlight the Turing's data science and AI expertise and desire to help the DRCF tackle the complex topics outlined in the plan of work for 2022 to 2023.

2.1 Algorithmic transparency

Algorithmic transparency, one of the priority areas in DRCF's plan of work for 2022-23 under the heading of collaboration, is where The Alan Turing Institute has deep pockets of expertise and foundational work. We would be keen to help support the DRCF's work in this important area. Algorithmic transparency requires careful thought and consideration. Transparency is a necessary but, on its own, is an insufficient aim in the attempt to monitor the design, development, and deployment of a particular AI system. For instance, an organisation can be "transparent" by making technical and complex information available to stakeholders who lack the requisite expertise to understand or interpret the information, or by overloading a regulator/auditor with logs that require a great deal of time to review. Therefore, careful consideration must be given to the specific mechanisms by which transparency will be ensured and enhanced through additional principles that secure the desired result.

Moreover, algorithmic transparency is a complex and multivalent concept that likely cannot be reduced to providing information about the underlying rationale of an AI model. Until recently, technology-centred understandings of algorithmic transparency dominated conversations about how to explain AI-assisted decisions and how to ensure justifiable outcomes for impacted data subjects. However, research led by the Turing has demonstrated the potential shortcomings of this perspective, showing instead that an interdisciplinary and practice-based approach to algorithmic transparency enables more effective technology policy and more responsible AI innovation. This richer understanding of transparency described above involves at least three distinctive levels:

- 1. Process transparency** requires that the design, development, and implementation processes underlying the decisions or behaviours of data systems are accessible for oversight and review, so that justified public trust and public consent can be ascertained.
- 2. Professional and institutional transparency** requires that, at every stage of the design and implementation of a project, responsible team members should be identified and held to rigorous standards of conduct that secure and maintain professionalism and institutional transparency. These standards should include the core, justice-promoting values of integrity, honesty, and sincerity, as well as positionality-aware modes of neutrality, objectivity, and impartiality. Professionals involved in the research, development, production, and implementation of data-intensive technologies are, first and foremost, acting as fiduciaries of the public interest and should, in keeping with these core justice-promoting values, put the obligations to serve that interest above any other concerns.
- 3. Outcome transparency** demands that stakeholders are informed of where data systems are being used and how and why such systems performed the way they did in specific contexts. Outcome transparency, therefore, requires that impacted individuals can understand the rationale behind the decisions or behaviours of these systems, so that they can contest objectionable results and seek effective remedy. Such information should be provided in a plain, understandable, non-specialist language and in a manner relevant and meaningful to those affected.

2.2 Relevant Turing Research

There are several existing pieces of work done by researchers within the Turing Institute Public Policy Programme, that are salient to this more nuanced approach to transparency:

- The report on ***AI in Financial Services***, commissioned by the Financial Conduct Authority, examines the fundamental role of AI transparency in pursuing responsible innovation. It shows how transparency can act as an enabler for other AI ethics principles such as fairness, sustainability, safety, and accountability, and argues that transparency is a logical first step for considering responsible AI innovation. Crucially, the report has a dedicated chapter on AI transparency, which outlines: the forms that AI transparency can take; the purposes it can serve; and relevant practical considerations for firms and regulators interested in AI transparency to demonstrate that AI systems are trustworthy and used responsibly.
- The joint guidance from the ICO and the Turing on ***Explaining decisions made with AI***, which sets out best practices for AI explainability. Here, the concept ‘explainability’ is used in addition to a related term ‘interpretability,’ which tends to emphasise technical properties of algorithmic systems (e.g. complexity of technique used). By contrast, ‘explainability’ draws attention to wider system design, development, and deployment processes that may be important for ensuring stakeholders understand why an algorithmic system is functioning in a particular way (i.e., focusing on both outcomes and the processes by which they arise). Constructing suitable or sufficient explanations for the behaviour of algorithmic systems places an ex-ante responsibility on the system’s developers to go beyond minimal forms of transparency, and promotes best practices that would support the priorities of the DRCF - this guidance has already been referenced by the DRCF in the Spring 2022 **publication on auditing algorithms**, and so the Turing would welcome the opportunity to work with the DRCF to find ways to integrate the guidance into the Forum’s proposed priorities.
- The methodology of argument-based assurance, which was introduced in Annex 5 of the joint guidance with the ICO on ***Explaining decisions made with AI***, and remains ongoing work in the Turing’s Public Policy Programme. As we in a more recent publication, argument-based assurance is a “process of using structured argumentation to provide reviewable (or, auditable) assurance that a particular set of claims about properties of an algorithmic system are warranted given the available evidence.” The method of argument-based assurance is already used in safety-critical domains, such as aviation, to support compliance practices. More recently, our work has extended the methodology in several ways that could be relevant to the DRCF. In particular, the methodology is anchored in a robust and systematic approach to delineating the key processes and activities that exist in a typical project lifecycle for an algorithmic system (see Figure 1). This model can help regulators by setting out a shared framework for identifying critical regulatory touchpoints, which are linked to important socioeconomic value chains.
- Furthermore, the updated assurance methodology that has been developed at the Turing is accompanied by an operationalisable set of principles and attributes that allow developers and regulators to understand relevant harms and benefits that can occur during the design, development, deployment, and use of algorithmic systems. We detail these principles and attributes in a **recent proposal to the Council of Europe’s** Ad hoc Committee on Artificial Intelligence. This proposal focused on the Human Rights impact of AI systems, but the salience of the principles and attributes also extends to much of the work conducted by DRCF’s four regulators (e.g., see Figure 2 on the principle of Accountability).

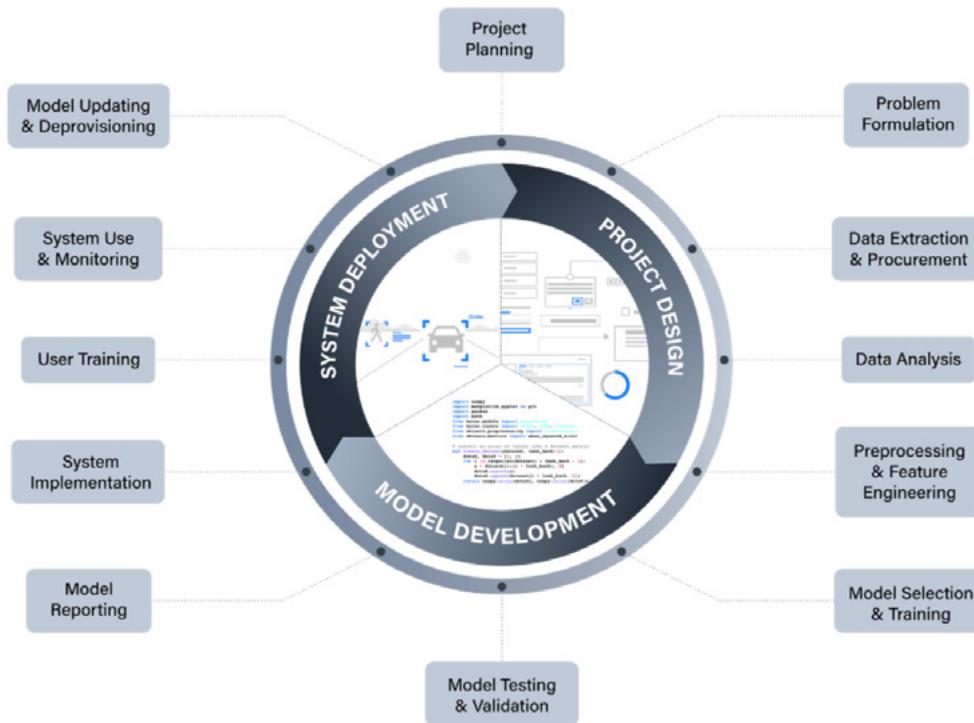


Figure 1: A model of the lifecycle for a typical project involving the design, development, and deployment of an algorithmic or data-driven system (reprinted from **Burr and Leslie, 2021**). Each stage (e.g., Project Design) and corresponding activity (e.g. Data Analysis) serves as a regulatory touchpoint for interventions that can generate additional socioeconomic value through processes of responsible research and innovation.

Accountability

Core Attribute	Description
Traceability	Traceability refers to the process by which all stages of the data lifecycle from collection to deployment to system updating or deprovisioning are documented in a way that is accessible and easily understood. This may include not only the parties within the organisation involved but also the actions taken at each stage that may impact the individuals who use the system.
Answerability	Answerability depends upon a human chain of responsibility. Answerability responds to the question of who is accountable for an automation supported outcome.
Auditability	Whereas the property of answerability responds to the question of who is accountable for an automation supported outcome, the notion of auditability answers the question of how the designers and implementers of AI systems are to be held accountable. This aspect of accountability has to do with demonstrating and evidencing both the responsibility of design and use practices and justifiability of outcomes.
Clear Data Provenance and Lineage	Clear provenance and data lineage consists of records that are accessible and simultaneously detail how data was collected and how it has been used and altered throughout the stages of pre-processing, modelling, training, testing and deploying.
Accessibility	Accessibility involves ensuring that information about the processes that took place to design, develop, and deploy an AI system are easily accessible by individuals. This not only refers to suitable means of explanation (clear, understandable, and accessible language) but also the mediums for delivery.
Reproducibility	Related to and dependant on the above four properties, reproducibility refers to the ability for others to reproduce the steps you have taken throughout your project to achieve the desired outcomes and where necessary to replicate the same outcomes by followings the same procedure.

Figure 2: An example of one of the principles introduced in **The Alan Turing Institute's proposal to the Council of Europe**, with the corresponding attributes that help specify and operationalise the principle in the context of an algorithmic system.

3. Capability

The Turing considers the section on capability the most important part of the DRCF's plan of work for 2022 to 2023. Capability is an enabler for the other two overarching goals, coherence and collaboration - without it, none of the other work of the DRCF would be possible.

Increasingly, regulators need access to cutting edge knowledge about the state of the art in AI. Access to expertise is essential for regulators if they are to anticipate future applications of AI as well as their respective benefits and risks. It is also crucial for ensuring that regulators can ask relevant and appropriate questions of regulated entities.

Simultaneously, regulators have significant opportunities to use AI technologies as regulatory tools. Developing and deploying AI tools will become essential for dealing with increasingly large volumes of data and regulating innovative uses of technology. This is particularly evident in the area of online safety, where regulation is required to address a wide range of harms (e.g. related to misinformation, disinformation, hate speech, or micro-targeting) and where AI has an important role to play as a regulatory tool (e.g. through automatic detection of problematic content, automatic fact-checking, hate speech detection tools, or systems to protect the digital privacy of children).

3.1 Building common capacity in AI and regulation

Recent research at the Turing found that many of the challenges that regulators face in relation to AI are common among regulators of all sizes and sectors. There is significant value in developing common and collaborative approaches to build regulators' capacities and establish cross-regulator approaches to regulation and AI. Regulators must collaborate to ensure consistent, complementary, and effective regulation.

Our research has also explored current capacities of UK regulatory bodies in relation to AI - both in terms of regulation of AI and of using AI systems as regulatory tools. It found that UK regulators were at varying levels of maturity and had different levels of capability in terms of both regulating and adopting AI. In particular, smaller regulators and those working in sectors without a strong history of work in relation to innovation/technology were least advanced in relation to their AI capabilities. Even larger regulators struggle to keep pace with technological regulation.

For example, in the area of online content moderation, end-2-end encrypted services are incredibly difficult to moderate, as the platform on which the messages are exchanged cannot access the content. Without moderation, platforms must rely on users reporting bad content, which is often ineffective against many sources of harm, such as child abuse, hate, or extremism, or they must use the metadata signals, such as the number of messages being sent in a given period of time. This can be useful for identifying scams, fraud, and misinformation, which have distinctive patterns of diffusion and spread online, but is ineffective at identifying most other sources of harm. Researchers are looking at on-device moderation for end-2-end encryption to solve this problem, but the technology is in its early stages. Given the small number of experts in this area, even a large, well-resourced regulator will likely struggle to truly assess how safety can be ensured with platforms that use end-2-end encryption. Similar challenges exist for regulating safety on services which use even more emergent forms of technology, such as blockchain-based platforms.

Setting aside the difficulties of building capacity and expertise, our recommendations for the DRCF are:

- **to ensure that the DRCF and its constituent members build capacities across all roles/teams.** As AI and digital innovation are increasingly relevant to all functions of an organisation, skills and capabilities should not be located only within particular departments or teams. Skills development and training opportunities for AI should be accessible to people at all levels of seniority and across all departments/teams. Moreover, AI and digital innovation are interdisciplinary fields, requiring expertise in technical, policy, ethical, legal, and governance dimensions – to name only a few. Therefore, developing regulatory capabilities needs to reach right across the organisation, embracing and developing diverse forms of expertise and skills.
- **to foster buy-in and understanding at senior management level.** Turing research has found that in regulatory bodies with the most advanced/mature AI capabilities, there was a strong commitment, understanding, and leadership from senior management. Such buy-in is necessary to underpin continuous investment and enable a degree of risk-taking (i.e. using new methods and approaches).
- **to develop mechanisms to ensure that learning and opportunities are shared beyond the DRCF members** in order to add value across the regulatory landscape. Regulation and AI/digital innovation challenges are common across regulators in all sectors and of all sizes. Yet smaller regulators have the least resources to draw on to develop new capabilities or establish internal expertise.
- **to offer structured or formal training and skills development regarding AI that is available to regulators.** Developing this area of work will be of value and necessity well beyond the members of the DRCF. The Turing is well-placed to help in this endeavour, as we are working on developing a skills programme with a strong focus on ethics and responsible innovation.

3.2 Standardisation

We also recommend that the DRCF consider the rapidly evolving field of AI standardisation as an important area of capability building. Standards developed in international Standards Development Organisations (SDOs) are set to play an influential role as governance and innovation mechanisms for AI. Recent years have seen a growing number of AI-related standards being developed by prominent SDOs including ISO, IEC, IEEE, and others. In the wake of the EU's proposed AI Act, the landscape of emerging AI-related standards will expand further as the European Standards Organisations (CEN, CENELEC, and ETSI) respond to the request to develop Harmonised Standards in relation to the EU AI Act. In parallel to these developments at the international level, the UK has seen the introduction of the concept of Designated Standards and recent government publications such as the [***Plan for Digital Regulation***](#) and the CDEI's [***Roadmap to an Effective AI Assurance Ecosystem***](#) have emphasised the role of standards in governing AI and other digital technologies.

These developments raise a wide range of novel questions for regulators whose remit touch on AI, especially in the case of organisations with a limited history of engaging with standardisation in other areas. These questions include:

- keeping track of the large amount of AI standards being developed and identifying those that are relevant to individual regulatory remit
- understanding standards development processes and, where appropriate, participating in them
- assessing the quality and adequacy of published standards, and
- making informed decisions about the relationship between standards and regulatory expectations.

Earlier this year, the Turing was selected to lead the pilot of the [AI Standards Hub](#), as set out in the Government's [National AI Strategy](#). Delivered in partnership with the British Standards Institution and the National Physical Laboratory, this new initiative aims to provide a platform to advance discussions about AI standardisation in the UK and equip stakeholders with the knowledge and tools to actively engage with the development and use of standards for AI. Addressing the needs of UK regulators in navigating AI standardisation is a key element of this mission and we would be delighted to collaborate with the DRCF with a view to enabling capability building in this area and promoting coherence between the evolution of standards and regulatory strategies in relation to AI.

3.3 Stakeholder inclusion and engagement

Apart from developing in-house skills and expertise and building capacity in the area of AI standardisation, we want to underline the importance of broadening stakeholder inclusion and engagement. Among the challenges of increasing regulator capacities around AI is for regulators to gain greater exposure to experiential expertise with technology; that is, digital policy deregulation often does not take the end-user experience and the social effects of technologies into account because of the limited means for acquiring knowledge of on-the-ground experience, including the role played by digital systems in the safety and well-being all communities.

The goal of promoting innovation is fully aligned with the responsibility of regulators to ensure a safe, sustainable, healthy, and democratic society. Many digital systems, including those involving AI, can often reproduce or exacerbate social and economic problems. The people most affected by these problems are the ones who can understand the problems most clearly. Gender, ethnic, and religious minorities, immigrants, refugees and those living with the least wealth are also the subjects of digital technologies and have important insights that qualify them as experts with standing in stakeholder consultations. Such experts should be counted as essential participants.

Where direct participation by these “experiential experts” can be accomplished, it must be conducted under conditions that make it meaningful. Just as regulators themselves benefit from increased expertise in AI and the other technologies they regulate, experiential experts can also effectively advocate for the interests of their communities only when they have opportunities to learn about how such systems work and how their lives are affected by them. We recommend that:

- **the DRCF prioritises investing in and building mechanisms to recruit and train experiential experts** and to ensure that their voices are at least as well-represented in regulator consultations as those of the technology industry and their supporters.
- **expert networks include experiential experts, namely ordinary people who can provide insights about the effects of AI and other digital technologies on their lives.** The DRCF can champion the creation of consultation groups drawn from the general population and integrate them into expert networks. Such networks should be developed in addition to the inclusion of members of representative organisations, such as civil society groups.
- **experiential experts are drawn from a broad range of life experiences.** To rebalance any historical patterns of exclusion in decision-making contexts, membership of such groups should consider having overrepresentation of people of minoritised groups and those who otherwise possess less social, economic, or political power than the stakeholders who are typically consulted.

- **capability building includes the upskilling of those providing experiential expertise.** The DRCF can pilot a training programme for members of the public in the fundamentals of AI and the digital technologies that are potential or existing targets for regulation. For example, many people are unaware that search engines and social media platforms personalise content for each user based on their perceived interests for the purpose of targeting advertising. While personalisation is sometimes welcome as a form of curation, it can often serve narrow business goals rather than wider societal interests. Consequently, the content, which may appear neutral and objective, can instead be designed to keep the user on the platform for revenue-generating reasons, even when this is accomplished with potentially divisive, misleading, and anti-social messaging. When those using such platforms understand this, they are more likely to seek safer and more prosocial options. By providing sufficient understanding of the technical design choices of information platforms, the users of those systems are better able to offer objections or ideas for their regulation and use.

4. Closing comments

The Turing welcomes the publication of the DRCF's plan of work for 2022 to 2023, particularly the focus on the overarching aims of coherence, collaboration, and capacity building. The work that the DRCF will undertake in these crucial areas will pave the way for the UK to lead the global conversation on the regulation of online environments and data-intensive technologies. As the national institute for data science and AI, the Turing stands ready to help the DRCF reach its important goals. Our expertise and in-depth research into the regulation of AI can be an invaluable resource to the DRCF as it embarks on its journey towards delivering on the ambition of the plan of work for 2022 to 2023.

5. Contributors

The following individuals (listed alphabetically) from the Turing have contributed to this response:

- Dr Mhairi Aitken, Ethics Research Fellow, Public Policy Programme, Turing Institute
- Morgan Briggs, Research Associate for Data Science and Ethics, Public Policy Programme, Turing Institute
- Dr Christopher Burr, Ethics Fellow, Public Policy Programme, Turing Institute
- Professor Jon Crowcroft, Turing, Researcher-at Large, Turing Institute; and University of Cambridge
- Dr Cosmina Dorobantu, Policy Fellow and Co-Director, Public Policy Programme, Turing Institute
- Dr Michael Katell, Ethics Fellow, Public Policy Programme, Turing Institute
- Dr David Leslie, Director of Ethics and Responsible Innovation Research, Public Policy Programme, Turing Institute
- Professor Helen Margetts, Turing Fellow and Director, Public Policy Programme, Turing Institute; and University of Oxford
- Dr Florian Ostmann, Head of AI Governance and Regulatory Innovation, Public Policy Programme, Turing Institute
- Dr Bertie Vidgen, Head of Online Safety, Public Policy Programme, Turing Institute

For comments or questions regarding this response, please contact policy@turing.ac.uk.



turing.ac.uk
@turinginst